

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification⁶ : C12N 15/11, 15/63, 15/85, 5/14, 15/00, 15/82		A1	(11) International Publication Number: WO 99/10488 (43) International Publication Date: 4 March 1999 (04.03.99)
(21) International Application Number: PCT/US98/17852 (22) International Filing Date: 28 August 1998 (28.08.98) (30) Priority Data: 08/919,501 28 August 1997 (28.08.97) US (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 08/919,501 (CIP) Filed on 28 August 1997 (28.08.97) (71) Applicant (for all designated States except US): THE SALK INSTITUTE FOR BIOLOGICAL STUDIES [US/US]; 10010 North Torrey Pines Road, La Jolla, CA 92037 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): O'GORMAN, Stephen [US/US]; 4647 Leathers Street, San Diego, CA 92117 (US). WAHL, Geoffrey [US/US]; 4258 Arquello Street, San Diego, CA 92103 (US). (74) Agent: REITER, Stephen, E.; Gray Cary Ware & Freidenrich LLP, Suite 1600, 4365 Executive Drive, San Diego, CA 92121 (US).			(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i>
(54) Title: SITE-SPECIFIC RECOMBINATION IN EUKARYOTES AND CONSTRUCTS USEFUL THEREFOR			
(57) Abstract <p>Site-specific recombinases provide a means of efficiently manipulating chromosomal sequences in mammalian cells in culture. Five lines of transgenic mice containing a fusion gene encoding the mouse protamine 1 gene promoter and the Cre recombinase gene, ProCre, showed high levels of Cre-mediated recombination in the germline, but did not show appreciable recombination in other tissues. In other transgenic mice containing a Cre targeting construct and the ProCre construct, between 80–100 % of the progeny inherited the Cre-recombined target. When ES cells from the ProCre line were transfected with vectors containing a loxP targeting construct, clones in which the loxP sites remained functional were readily isolated. These data establish that ProCre nucleic acid constructs will facilitate the production of subtle, conditional or tissue-specific mutations in mice as well as the production and analysis of mice with recombinase-conditional lethal alleles.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Site-Specific Recombination in Eukaryotes
and Constructs Useful Therefor

FIELD OF THE INVENTION

The present invention relates to methods for manipulating chromosomal sequences in cells by site-specific recombination promoted by recombinases. In a particular aspect, the present invention relates to methods for producing embryonic stem cells bearing nucleic acid sequences that have been rearranged by a site-specific recombinase expressed from a construct controlled by a tissue-specific promoter (e.g., a germline specific promoter). In another aspect, the present invention relates to methods for producing embryonic stem cells bearing nucleic acid sequences that have been rearranged by a site-specific recombinase expressed from a construct controlled by a conditional promoter.

BACKGROUND OF THE INVENTION

The analysis of gene function has increasingly come to require the production of subtle, tissue-specific, and conditional mutations in animals and plants. Although there are a number of methods for engineering subtle mutations in embryonic stem (ES) cells (Hasty et al. (1991) *Nature* 350:243-246, Askew et al. (1993) *Mol Cell Biol* 13:4115-4124), the use of site-specific recombinases to remove the selectable marker that permits isolation of homologously recombined ES cell clones has become increasingly prevalent (Kitamoto et al. (1996) *Biochem Biophys Res Commun* 222:742-747, Fiering et al. (1993) *Proc Natl Acad Sci U S A* 90:8469-8473, Schwenk et al. (1995) *Nucleic Acids Res* 23:5080-5081; Gu et al. (1993) *Cell* 73:1155-1164; Sailer et al. (1996) *Taniguchi Symposia on*

Brain Sciences, eds. Nakanishi et al. (Japan Scientific Press), pp. 89-98).

Site-specific recombinases represent the best method for creating tissue-specific and conditional mutations in animals and plants, being employed first to remove the selectable marker to create a functionally wild-type allele, and then to inactivate the allele mosaically in animals and plants by removing some essential component in a tissue-specific or conditional manner (Gu et al. (1994) *Science* 265:103-106; Kuhn et al. (1995) *Science* 269:1427-1429). Current protocols for using excissive site-specific recombination to remove selectable markers include transiently transfecting ES cell clones with a recombinase expression vector (Gu et al. (1993) *Cell* 73:1155-1164), microinjecting fertilized oocytes containing the recombinant allele with a recombinase expression vector (Kitamoto et al. (1996) *Biochem Biophys Res Commun* 222:742-747; Araki et al. (1995) *Proc Natl Acad Sci U S A* 92:160-164), or breeding animals and plants containing the recombinant allele to animals and plants, respectively, containing a recombinase transgene (Schwenk et al. (1995) *Nucleic Acids Res* 23:5080-5081; Lewandoski et al. (1997) *Curr Biol* 7:148-151). Each of these approaches requires an investment of some combination of time, resources, and expertise over that required to generate animals and plants with homologously recombined alleles. The most commonly employed method, the secondary transfection of homologously recombined ES cell clones with a recombinase expression vector, additionally requires extended culture time that may decrease their potential to enter the germline.

In principle, marker excision would be substantially simplified through the use of ES cells containing

recombinase nucleic acid constructs that were expressed in the germline, but not to an appreciable extent in the ES cells themselves or somatic tissues of animals and plants. The lack of ES cell expression would mean that targeting
5 vectors containing selectable markers flanked by recombinase target sites could be used to isolate homologous recombinants without fear that the marker would be excised during culture. Robust recombinase expression in gametes would mean that the marker would be excised in
10 at least some of the progeny of ES cell chimeras. Only a single step would be required to isolate subtle mutations and, if two different recombinase systems were employed, conditional and tissue-specific alleles could be produced with similar improvements in efficiency. A
15 germline-specific recombinase nucleic acid construct could also be used to deliver recombined target nucleic acid constructs to the early embryo (Lewandoski et al. (1997) *Curr Biol* 7:148-151), so long as the recombined target was not detrimental to the terminal stages of spermatogenesis.

20 Previous reports have shown that expression of nucleic acid constructs containing the proximal promoter of the mouse protamine 1 (mP1) locus is restricted to haploid spermatids in mature mice (Peschon et al. (1987) *Proc Natl Acad Sci U S A* 84:5316-5319; Behringer et al. (1988) *Proc*
25 *Natl Acad Sci U S A* 85:2648-2652), although low levels of ectopic expression may occur in some mature tissues (Behringer et al. (1988) *Proc Natl Acad Sci USA* 85:2648-2652). Inclusion of the mP1 promoter does not guarantee expression in the male germline, however, for
30 although nucleic acid constructs containing the mP1 promoter and the SV40 T-antigen coding sequence were transcribed, the message was not translated at detectable

levels in spermatids (Behringer et al. (1988) *Proc Natl Acad Sci USA* 85:2648-2652).

Accordingly, there is a need in the art for methods to modulate expression of recombined target nucleic acid sequences in the early embryo. In addition, there is a need in the art for tissue-specific and conditional recombinatory tools to create transgenic animals and plants. These and other needs in the art are addressed by the present invention.

10 BRIEF DESCRIPTION OF THE INVENTION

The present invention meets the need in the art for modulating expression of recombined target nucleic acid sequences to the early embryo. The present invention further meets the need in the art for tissue-specific and conditional recombinatory tools to create transgenic animals and plants. Thus, in accordance with the present invention, it has been discovered that nucleic acid constructs encoding a germline specific promoter operatively associated with a recombinase coding sequence lead to efficient recombination of a target nucleic acid construct in the male germline, but not in other tissues. This suggests that such nucleic acid constructs could be used for the efficient production of embryos bearing conditional, genetically lethal alleles. It has additionally been discovered that ES cell lines generated from one of these transgenic lines could be used in combination with targeting vectors that contained loxP-flanked selectable markers to isolate homologous recombinants containing the marker and functional loxP sites.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 illustrates a schematic of P2Bc and P2Br alleles. The positions of the PCR primers used (5'P and 3'P) are indicated on the diagrams of the P2Bc and P2Br alleles.

Figure 2 depicts the targeting of the *hoxb-1* locus in ProCre ES cells using a targeting vector that contains a loxP-flanked selectable marker. Top, schematic of the wild-type *hoxb-1* locus showing the positions of the two exons (open boxes), the position of a 5' NruI site and flanking BamHI restriction endonuclease sites, and PCR primers (triangles) that amplify a 204 bp product from the wild-type allele that includes the NruI site. Middle, the predicted organization of homologously recombined *hoxb-1* allele in which a neomycin cassette (NEO), flanked by loxP sites (L), has been inserted into the NruI site shown in the top diagram. The insertion creates a novel BamHI site and the same PCR primers now amplify a 1600 bp product. Bottom: the predicted structure of the recombined allele shown in the middle panel after Cre-mediated excision of the neomycin cassette to leave a single loxP site in place of the NruI site of the wild-type allele. Amplification with the same primers now yields a 268 bp product.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the present invention, there are provided nucleic acid constructs comprising a germline-specific promoter operatively associated with a recombinase coding sequence.

As used herein, the term "promoter" refers to a specific nucleotide sequence recognized by RNA polymerase, the enzyme that initiates RNA synthesis. The promoter sequence is the site at which transcription can be specifically initiated under proper conditions. The recombinase nucleic acid(s), operatively linked to the suitable promoter, is(are) introduced into the cells of a suitable host, wherein expression of the recombinase nucleic acid(s) is(are) controlled by the promoter.

10 Germline-specific promoters contemplated for use in the practice of the present invention include the protamine 1 gene promoter, the protamine 2 gene promoter, the spermatid-specific promoter from the c-kit gene (Albanesi et al. (1996) *Development* 122(4):1291-1302), the sperm-
15 specific promoter from angiotensin-converting enzyme (Howard et al. (1993) *Mol Cell Biol* 13(1):18-27; Zhou et al. (1995) *Dev Genet* 16(2):201-209), oocyte specific promoter from the ZP1 gene, oocyte specific promoter from the ZP2 gene, oocyte specific promoter from the ZP3 gene
20 (Schickler et al. (1992) *Mol Cell Biol* 12(1):120-127), and the like.

In addition to the above-described germline-specific promoters, tissue-specific promoters specific to plants are also contemplated for use in the practice of the present
25 invention, including, for example, the LAT52 gene promoter from tomato, the LAT56 gene promoter from tomato, the LAT59 gene promoter from tomato Eyal et al. (1995) *Plant Cell* 7(3):373-384), the pollen-specific promoter of the Brassica S locus glycoprotein gene (Dzelzkalns et al. (1993) *Plant*
30 *Cell* 5(8):855-863), the pollen-specific promoter of the NTP303 gene (Weterings et al. (1995) *Plant J* 8(1):55-63), and the like.

Recombinases contemplated for use in the practice of the present invention include Cre recombinase, FLP recombinase, the R gene product of *Zygosaccharomyces* (Onouchi et al. (1995) *Mol Gen Genet* 247(6):653-660), and
5 the like.

Presently preferred constructs contemplated for use in the practice of the present invention include ProCre (comprising the protamine 1 gene promoter operatively associated with Cre recombinase), ProFLP (comprising the
10 protamine 1 gene promoter operatively associated with FLP recombinase), ProR (comprising the protamine 1 gene promoter operatively associated with the R gene product of *Zygosaccharomyces*), and the like.

In accordance with another embodiment of the present
15 invention, there are provided nucleic acid constructs comprising a conditional promoter or a tissue-specific promoter operatively associated with a recombinase coding sequence.

Promoters contemplated for control of expression of
20 recombinase nucleic acid(s) employed in accordance with this aspect of the present invention include inducible (e.g., minimal CMV promoter, minimal TK promoter, modified MMLV LTR), constitutive (e.g., chicken β -actin promoter, MMLV LTR (non-modified), DHFR), and/or tissue specific
25 promoters.

Conditional promoters contemplated for use in the practice of the present invention comprise transcription regulatory regions that function maximally to promote transcription of mRNA under inducing conditions. Examples
30 of suitable inducible promoters include DNA sequences

corresponding to: the *E. coli* lac operator responsive to IPTG (see Nakamura et al., *Cell*, 18:1109-1117, 1979); the metallothionein promoter metal-regulatory-elements responsive to heavy-metal (e.g., zinc) induction (see Evans et al., U.S. Patent No. 4,870,009), the phage T7lac promoter responsive to IPTG (see Studier et al., *Meth. Enzymol.*, 185: 60-89, 1990; and U.S. #4,952,496), the heat-shock promoter; the TK minimal promoter; the CMV minimal promoter; a synthetic promoter; and the like.

10 Exemplary constitutive promoters contemplated for use in the practice of the present invention include the CMV promoter, the SV40 promoter, the DHFR promoter, the mouse mammary tumor virus (MMTV) steroid-inducible promoter, Moloney murine leukemia virus (MMLV) promoter, elongation
15 factor 1 α (EF1 α) promoter, albumin promoter, APO A1 promoter, cyclic AMP dependent kinase II (CaMKII) promoter, keratin promoter, CD3 promoter, immunoglobulin light or heavy chain promoters, neurofilament promoter, neuron specific enolase promoter, L7 promoter, CD2 promoter,
20 myosin light chain kinase promoter, HOX gene promoter, thymidine kinase (TK) promoter, RNA Pol II promoter, MYOD promoter, MYF5 promoter, phosphoglycerokinase (PGK) promoter, Stf1 promoter, Low Density Lipoprotein (LDL) promoter, chicken β -actin promoter (used in conjunction
25 with ecdysone response element) and the like.

As readily understood by those of skill in the art, the term "tissue specific" refers to the substantially exclusive initiation of transcription in the tissue from which a particular promoter, which drives expression of a
30 given gene, is derived (e.g., expressed only in T-cells, endothelial cells, smooth muscle cells, and the like). Exemplary tissue specific promoters contemplated for use in

the practice of the present invention include the GH promoter, the NSE promoter, the GFAP promoter, neurotransmitter promoters (e.g., tyrosine hydroxylase, TH, choline acetyltransferase, ChAT, and the like), promoters
5 for neurotropic factors (e.g., a nerve growth factor promoter, NT-3, BDNF promoters, and the like), and so on.

In accordance with yet another embodiment of the present invention, there are provided embryonic stem cells containing a nucleic acid construct as described herein.

10 As readily understood by those of skill in the art, the above-described constructs can be introduced into a variety of animal species, such as, for example, mouse, rat, rabbits, swine, ruminants (sheep, goats and cattle), humans, poultry, fish, and the like. Transgenic
15 amphibians, insects, nematodes, and the like, are also contemplated. Members of the plant kingdom, such as, for example, transgenic mono- and dicotyledonous species, including important crop plants, i.e., wheat, rice, maize, soybean, potato, cotton, alfalfa, and the like, are also
20 contemplated.

For example, pluripotential ES cells can be derived from early pre-implantation embryos, preferably the ova are harvested between the eight-cell and blastocyst stages. ES
25 cells are maintained in culture long enough to permit integration of the promoter-recombinase nucleic acid construct(s). The cells are then either injected into a host blastocyst, i.e., the blastocoel of the host blastocyst, or co-cultured with eight-cell to morula-stage
30 ova, i.e., zona-free morula, so that transfected ES cells are preferentially incorporated into the inner cell mass of the developing embryo. With blastocyst injection,

transgenic offspring are termed "chimeric," as some of their cells are derived from the host blastocyst and some transfected ES cells. The host embryos are transferred into intermediate hosts or surrogate females for continuous
5 development.

The transformation procedure for plants usually relies on the transfer of a transgene carrying a particular promoter construct via the soil bacterium *Agrobacterium tumefaciens*. Transformation vectors for this procedure are
10 derived from the T-DNA of *A. tumefaciens*, and transgenes are stably incorporated into the nuclear genome. The activity of the transgenes can then be monitored in the regenerated plants under different conditions. In this way, many promoter elements that are involved in complex
15 regulatory pathways such as light responsiveness or tissue specificity have been defined.

Alternatively, direct (i.e., vectorless) gene transfer systems are also contemplated including chemical methods, electroporation, microinjection, biolistics, and
20 the like. Protoplasts isolated from the plants can be obtained by treatment with cell wall degrading enzymes. DNA can be introduced into plant protoplasts by a number of physical techniques including electroporation and polyethylene glycol treatment in the presence of $MgCl_2$.
25 The method of choice for rapid promoter analyses in plants is the biolistic method. This technique involves the delivery of the particular DNA construct into plant cells by microprojectiles, i.e., nucleic acid(s) coated or precipitated by tungsten or gold. This method is not
30 limited to any particular plant species or tissue type. Preferably, this method would allow quantitative analysis

of transformation if appropriate selectable markers are included.

In a preferred embodiment, the genome of embryonic stem cells according to the invention comprise a transcriptionally active selectable marker flanked by two recombination target sites. It is especially preferred that the recombinase encoded by the recombinase coding sequence operatively associated with a germline-specific promoter is selective for the recombination target sites flanking said selectable marker.

Optionally, embryonic stem cells according to the invention may further comprise one or more of:

a nucleic acid fragment flanked by two recombination target sites, wherein said recombination target sites are different than the recombination target sites which flank said selectable marker,

a nucleic acid construct comprising a conditional promoter operatively associated with a recombinase coding sequence,

a second nucleic acid construct comprising a tissue-specific promoter operatively associated with a second recombinase coding sequence, or the like. Preferably, the second recombinase coding sequence will be different than the first recombinase coding sequence.

The ability to select and maintain nucleic acid constructs in the host cell is an important aspect of an expression system. The most common type of selectable marker incorporated in the nucleic acid construct is an antibiotic resistance element allowing selection with ampicillin, kanamycin, neomycin, tetracycline, hygromycin, puromycin, blastophycin, and the like. Other approaches

employ specially constructed host cells which require the selectable marker for survival. Such selectable markers include the valine tRNA synthetase, val^S, the single-stranded DNA binding protein, *ssb*, thymidine kinase, or the like. Alternatively, naturally occurring partition systems that maintain copy number and select against plasmid loss is also contemplated. An example is the incorporation of the *parB* locus. Other selectable markers include HPRT and the like.

10 Selectable markers specific for plants include, the *gus A* (*uid A*), the *bar* gene, phosphinothricin and the like.

 In accordance with still another embodiment of the present invention, there are provided methods for excission of the transcriptionally active selectable marker from the above-described embryonic stem cells, said method comprising:

15 passaging the genome derived from said embryonic stem cells through gametogenesis (i.e., spermatogenesis or oogenesis).

20 Excission of marker as contemplated herein can cause a variety of end results, e.g., deletion of the marker or a nucleic acid sequence, gain of function or loss of function, replacement of function, and the like, as well as modulation of any one or more of these results.

25 Functions which are contemplated to be manipulated include regulating body size and growth rate, including recombining gene constructs which contain various growth hormone gene sequences. Other productivity traits that are targets include altering the properties or proportions of caseins, lactose, or butterfat in milk, increased

30

resistance to viral and bacterial diseases (i.e., "constitutive immunity" or germ-line transmission of specific, recombined antibody genes), more efficient wool production, and the like. Other functions which are contemplated to be modulated include development of lines of transgenic animals and plants for use in directing expression of transgenes encoding biologically active human proteins.

Agronomic traits which are contemplated to be modulated by use of the present invention include tolerance to biotic and abiotic stress, increased resistance to herbicides, pest damage, and viral, bacterial, and fungal diseases, improvement of crop quality (i.e., increase in nutritional value of food and feed), reduction of post-harvest losses, improvement of suitability and enlargement of the spectrum for processing (i.e., altered quantity and composition of endogenous properties, production of new compounds of plant or non-plant origin such as biopolymers or pharmaceutical substances).

In accordance with a still further embodiment of the present invention, there are provided methods for the production of recombinant alleles, said method comprising:

introducing a nucleic acid fragment flanked by at least two recombination target sites into embryonic stem cells as described herein, and

passaging the genome derived from said embryonic stem cells through gametogenesis.

As readily recognized by those of skill in the art, nucleic acid fragments can be introduced into ES cells by a variety of techniques, e.g., by homologous recombination,

random insertion, retroviral insertion, site specific-mediated recombination, and the like.

Nucleic acid fragments contemplated for use herein include fragments containing an essential portion of a gene
5 of interest.

In accordance with yet another embodiment of the present invention, there are provided methods for the production of recombinant alleles, said method comprising:

10 introducing at least one recombinase responsive construct
into embryonic stem cells as described herein,
wherein said construct(s) comprise(s) a nucleic
acid fragment and a selectable marker,
wherein said selectable marker is flanked by a
first pair of recombination target sites, and
15 wherein said nucleic acid fragment is flanked by
a second pair of recombination target sites,

passaging the genome derived from said embryonic stem cells
through gametogenesis.

In a presently preferred aspect, the first pair of
20 recombination target sites is recognized by a recombinase
which is expressed under the control of a germline-specific
promoter and said second pair of recombination target sites
is recognized by a recombinase which is expressed under the
control of a conditional promoter or a tissue specific
25 promoter.

Optionally, the embryonic stem cells employed herein
can further comprise a second nucleic acid construct
selected from constructs comprising a conditional promoter

operatively associated with a recombinase coding sequence, a construct comprising a tissue-specific promoter operatively associated with a recombinase coding sequence, and the like.

5 In accordance with still another embodiment of the present invention, there are provided methods for the conditional assembly of functional gene(s) for expression in eukaryotic cells by recombination of individual inactive gene segments from one or more gene(s) of interest,

10 wherein each of said segments contains at least one recombination target site, and

wherein at least one of said segments contains at least two recombination target sites,

said method comprising:

15 introducing said individual inactive gene segments into an embryonic stem cell as described herein, thereby providing a DNA which encodes a functional gene of interest, the expression product of which is biologically active, upon passage of the genome derived from said stem cells through gametogenesis.

20

For assembly of functional genes from inactive gene segments, see, for example, US Patent No. 5,654,182, incorporated herein by reference in its entirety.

25 In accordance with a still further embodiment of the present invention, there are provided methods for the generation of recombinant livestock, said method comprising:

combining embryonic stem cells that include nucleic acid construct according to the invention with host

30

pluripotential ES cells derived from early pre-implantation embryos, and

introducing these combined embryos into a host female and allowing the derived embryos to come to term.

5

In accordance with yet another embodiment of the present invention, there are provided methods for the generation of recombinant plants, said method comprising transforming plant zygotes with nucleic acid constructs
10 according to the invention and allowing the zygote to develop.

The objective of the current work with ProCre nucleic acid constructs was to determine the potential of germline-specific promoters to implement efficient
15 approaches utilizing site-specific recombinases to generate an array of sophisticated mutations in mammals and plants. The data shows that it is possible to create recombinase nucleic acid constructs that are expressed at high levels in the germ line but not to a functionally significant
20 extent in either ES cells or embryonic or adult somatic tissues. Homologous recombinants with a selectable marker can be isolated in ES cells that contain promoter-recombinase nucleic acid constructs. Transgenic animals and plants bearing the promoter-recombinase nucleic
25 acid constructs and a target allele transmit the recombined target to their progeny at high frequencies. These results establish the principle that mammals and plants containing loci that have been homologously recombined and then subsequently site-specifically recombined can be generated
30 simply by using ES cells with a suitable recombinase nucleic acid constructs for the initial targeting. By this mechanism, alleles containing a single recombinase target site and a mutation of interest can be produced in the

progeny of ES cell chimeras without any investment of time, expertise, or resources over that required to create an allele that still contains a selectable marker. The paradigm has obvious utility in the production of subtle and conditional mutations that require generation of alleles with minimal structural alterations. Because the presence and transcriptional activity of selectable markers can contribute to phenotypes in an unanticipated and unwanted manner (Fiering et al. (1995) *Genes Dev* 9:2203-2213); Olson et al. (1996) *Cell* 85:1-4), the approach will also be useful for generating null alleles.

Expression of the endogenous mPl locus (Hecht et al. (1986) *Exp Cell Res* 164:183-190), and mPl-driven nucleic acid constructs (Behringer et al. (1988) *Proc Natl Acad Sci U S A* 85:2648-2652; Braun et al. (1989) *Nature* 337:373-376; Zambrowicz et al. (1993) *Proc Natl Acad Sci U S A* 90:5071-5075) is restricted to haploid spermatids. Expression of mPl nucleic acid construct expression typically begins at haploid stages, and both RNA (Caldwell and Handel (1991) *Proc Natl Acad Sci U S A* 88:2407-2411) and proteins (Braun et al. (1989) *Nature* 337:373-376) diffuse through the spermatogenic syncytium. The result is a highly efficient recombination of target alleles and the segregation of recombinase and target nucleic acid constructs in the first generation.

Cre-mediated recombination proved to be highly testis-specific in ProCre mice. It is clear that the nucleic acid constructs are not expressed in the inner cell mass or in other early embryonic tissues. Cells from pre-implantation embryos intermingle extensively and the embryo as a whole is derived from a small number of cells (Beddington et al. (1989) *Development* 106:37-46; Soriano and Jaenisch (1986)

Cell 46:19-29). If ProCre nucleic acid constructs recombined target sequences during pre-implantation stages, at least a few percent of the cells in many tissues would contain the P2Br allele and Southern and PCR analyses showed that this was not the case. The ectopic Cre activity seen in some ProCre strains probably resulted from low levels of recombinase expression in later embryos or mature tissues, a finding consistent with the expression patterns of other mPl-driven nucleic acid constructs. Northern analyses have failed to reveal the expression of mPl-containing nucleic acid constructs in a variety of mature tissues (Peschon et al. (1987) *Proc Natl Acad Sci USA* 84:5316-5319; Behringer et al. (1988) *Proc Natl Acad Sci U S A* 85:2648-2652; Peschon et al. (1989) *Ann N Y Acad Sci* 564:186-197; Zambrowicz et al. (1993) *Proc Natl Acad Sci U S A* 90:5071-5075), but nucleic acid constructs containing the mPl promoter and the SV40 T-antigen led to the consistent development of tumors of the petrosal bone and right cardiac atrium (Behringer et al. (1988) *Proc Natl Acad Sci U S A* 85:2648-2652).

PCR assays represent a very sensitive assay for whether sufficient levels of Cre protein were produced to effect recombination. Importantly, they measured the cumulative level of recombination, for events that occurred at any stage of development are likely to have been propagated to, and might be amplified in, descendant populations. The highest level of ectopic recombination was that observed in cardiac ventricular tissue of strain which generated a signal approximately equivalent to that expected if the ratio between recombined and unrecombined alleles were 1:104. The activities observed in other strains were considerably lower than this, and one strain did not show any ectopic activity. None of the

strains showed evidence of recombination in the cardiac atria and the petrosal bone was not examined. These assays did not rule out the possibility that higher levels of recombination occur in tissues that were not examined or
5 that the low levels of recombination observed in some tissues reflected high levels of recombination in some component cell population.

These low levels of ectopic activity suggest that mpl-driven recombinase nucleic acid constructs could be
10 used for the production of embryos containing genetically lethal alleles. Some alleles created by homologous recombination in ES cells will prove to be lethal in heterozygotes, as was the case for an mRNA editing mutation of the GluR2 glutamate receptor subunit (Brusa et al.
15 (1995) *Science* 270:1677-1680). Germline transmission would be restricted to rare chimeras in which the level of chimerism was low enough in tissues affected by the mutation to allow survival and high enough in the germline to allow transmission. This problem could be circumvented
20 by creating recombinase-conditional mutations in ES cells bearing mpl-recombinase nucleic acid constructs, or by making the same mutations in standard ES cells and then introducing the mpl-recombinase nucleic acid construct by breeding. So long as the recombined version of the allele
25 did not adversely impact terminal stages of spermatogenesis, embryos containing the recombined allele could be efficiently produced. Embryos containing recombined nucleic acid constructs can also be produced through the activity of Cre nucleic acid constructs that
30 are expressed during early embryogenesis from the human cytomegalovirus minimal promoter (Schwenk et al. (1995) *Nucleic Acids Res* 23:5080-5081), the adenovirus E1A promoter (Lakso et al. (1992) *Proc Natl Acad Sci U S A*

89:6232-6236), or the zP3 promoter (Lewandoski et al. (1997) *Curr Biol* 7:148-151). ProCre and zP3 nucleic acid constructs have the advantage of delivering a recombined allele to the zygote, guaranteeing that all cells in the
5 derived embryos will contain the allele.

ProCre ES cells are but one of many different kinds of recombinase-bearing ES cells that could significantly shorten the time and effort required for a wide variety of genetic manipulations in mice. The most obvious of these
10 are complementary ProFLP ES cells in which the FLP recombinase was derived from *S. cerevisiae* (Broach and Hicks (1980) *Cell* 21:501-508) or another species (Kuhn et al. (1995) *Science* 269:1427-1429). Conceptually distinct from these but perhaps as generically useful would be ES cells
15 bearing inducible recombinase nucleic acid constructs that would facilitate temporal control of recombinase expression in ES cells, chimeras, and their progeny to generate site-specifically recombined alleles (Araki et al. (1992) *J Mol Biol* 225:25-37; No et al. (1996) *Proc Natl Acad Sci U S A* 93:3346-3351; Logie and Stewart (1995) *Proc Natl Acad Sci U S A* 92:5940-5944; Feil et al. (1996) *Proc Natl Acad Sci U S A* 93:10887-10890). Finally, fusion genes that led
20 to recombinase expression in specific tissues could be used to address specific research objectives.

25 The invention will now be described in greater detail by reference to the following non-limiting examples.

Example 1

Mammalian DNA Constructs

A 652 bp fragment of the mP1 promoter (SEQ ID NO:1;
30 Peschon et al. (1989) *Annals of the New York Academy of*

Sciences 186-197) was isolated by PCR using PCR primers (SEQ ID NOS:2 and 3) and genomic DNA templates from CCE ES cells (Robertson et al. (1986) *Nature* 323:445-448). This fragment was fused to a modified Cre coding sequence (SEQ ID NO:4) which contains a consensus translation start site (Kozak (1986) *Cell* 44:283-292), 11 codons for a human c-myc epitope (Evan et al. (1985) *Mol Cell Biol* 5:3610-3616), 7 codons for a minimal SV40 nuclear localization signal (Kalderon et al. (1984) *Cell* 39:499-509) and the polyadenylation signal from pIC-Cre in the plasmid pOG304M (SEQ ID NO:5). The Cre expression plasmid pOG231 was prepared by fusing a modified Cre coding sequence from pIC-Cre (Gu et al. (1993) *Cell* 73:1155-1164), and containing the same translation start and nuclear localization signal, to the synthetic intron and CMV promoter of pOG44 (O'Gorman et al. (1991) *Science* 251:1351-1355).

A plasmid, pOG277 (SEQ ID NO:7), containing a loxP-flanked neomycin cassette was prepared by inserting a wild-type loxP site (SEQ ID NO:8; Hoess et al. (1982) *Proc Natl Acad Sci U S A* 79:3398-402) into pBSKS (Stratagene) and then cloning the neomycin expression cassette from pMCIneo-polyA (Thomas et al. (1987) *Cell* 51:503-512) between interactions of this loxP site. The *hoxb-1* targeting construct consisted of the PGK-TK cassette from pPNT (Tybulewicz et al. (1991) *Cell* 65:1153-63), and 1.4kb and 10.2kb of sequences 5' and 3' to an *Nru I* site 800 bp 5' to the *hoxb-1* transcriptional start site isolated from a 129 strain genomic library (Stratagene). The loxP-flanked neo cassette from pOG277 was inserted into the *Nru I* site. The pOG277 neomycin cassette and a β -GAL sequence was inserted into the first exon of the large subunit of RNA polymerase II (RP2) (Ahearn et al. (1987)

J. Biol. Chem. 262:10695-10705) to create the P2Bc allele (Figure 1). Cre-mediated recombination of the P2Bc allele results in the deletion of the neomycin cassette (Neo) of P2Bc that is flanked by two loxP sites, leaving a single loxP site and fusing the B-Gal coding sequence to the initial codons of the RNA polymerase II coding sequence. Recombination increases the size of a Pst I fragment recognized by the RP2 probe, which is external to the targeting vector used, indicated by the shaded box below each allele.

Example 2

Production of transgenic mice

Fertilized oocytes obtained from matings of 129/SvJae (Simpson et al. (1997) *Nat Genet* 16:19-27) and BALB/c X C57BL/6 F1 mice were used for pronuclear injections of the Protamine-Cre fusion gene from pOG304M according to standard protocols (Hogan et al. *Manipulating the Mouse Embryo: The Manual*, Coldspring Harbor Press (1994), pg. 497). Production of ES cells and homologous recombinants: Heterozygous ProCre 129/SvJae males were mated to 129/SvEms-⁺_{Ter?}/J females (Simpson et al. (1997) *Nat Genet* 16:19-27) to produce blastocysts that were cultured according to standard protocols (Robertson (1987) *Teratocarcinomas and embryonic stem cells*, a practical approach, eds. E. J. Robertson (IRL Press), pp. 71-112). The sex (King et al. (1994) *Genomics* 24:159-68) and ProCre status of each line were determined by PCR assays. Molecular analyses: Tail biopsy genomic DNA was used for hybridization assays or PCR assays to identify ProCre and P2Bc/r mice. PCR reactions for the detection of ectopic Cre activity used 100 ng of genomic DNA as a template to

amplify a P2Br-specific product using a 5' primer from the RP2 promoter and a 3' primer from the β -GAL coding sequence (Figure 1). Thirty cycles of amplification were done in a total volume of 100 μ l using 300 ng of each primer, 3 mM MgCl₂, 1.5 units of Taq polymerase, and an annealing temperature of 60°C. Southern blots of reaction products were hybridized with a probe specific for the P2Br reaction product.

Example 3

ProCre Nucleic Acid Constructs Efficiently Recombine Target Alleles

A total of nine founder animals with ProCre nucleic acid constructs were obtained from injections of a Protamine-Cre fusion gene. Two lines were derived from injections of 129SvJae (Simpson et al. (1997) Nat Genet 16:19-27) embryos, and seven from injections of CB6F2 embryos. The 129/SvJae lines and three randomly selected hybrid lines were examined in detail. To determine whether ProCre nucleic acid constructs would efficiently recombine a target allele, males were generated that contained a ProCre nucleic acid construct and a target for Cre-mediated recombination. This "P2Bc" (Pol II, β -GAL, conditional) target (Figure 1) was created using homologous recombination in ES cells to insert a loxP-flanked neomycin cassette and a β -GAL coding sequence into the first exon of the locus coding for the large subunit of RNA polymerase II. Cre-mediated recombination of the loxP sites was expected to delete the intercalated sequences, creating "P2Br" allele (Pol II, β -Gal, recombined).

These males were mated to wild-type females and the resulting progeny were examined by Southern blotting to

determine if they inherited the P2Bc or the P2Br allele, and to additionally determine the segregation pattern of ProCre nucleic acid constructs and P2Br alleles. Southern blot of Pst I digested tail biopsy DNA's from a +/P2Bc, 5 +/ProCre male (sire) and four of his progeny by a wild-type female probed with n RP2 probe (top) and then reprobed with a Cre probe (bottom). The large majority of transmitted target alleles were Cre-recombined P2Br alleles (Table 1). ProCre nucleic acid constructs and recombined target 10 alleles segregated independently in the first generation; approximately 50% of mice that inherited a P2Br allele also inherited their male parent's ProCre nucleic acid construct. All RP2 mutant alleles in the progeny were P2Br, and some progeny inherit a P2Br allele without 15 inheriting ProCre nucleic acid construct. Mouse 4 did not contain a ProCre nucleic acid construct and is homozygous wild-type at the RP2 locus. These data establish that ProCre nucleic acid constructs efficiently recombine the P2Bc allele in the male germline and that the recombined 20 P2Br alleles and ProCre nucleic acid constructs segregate in the first generation. Because significantly more than 25% of the progeny inherited recombined target alleles, recombination either occurred during diploid stages of spermatogenesis or Cre generated during haploid stages of 25 spermatogenesis was distributed among spermatids through cytoplasmic bridges (Braun et al. (1989) Nature 337:373-376), effecting recombination in spermatids that did not themselves contain a ProCre nucleic acid construct.

The progeny of matings between ProCre males and +/P2Bc 30 females were also examined to determine if male gametes from ProCre mice delivered enough Cre to zygotes to effect Cre-mediated recombination of a target sequence. Of 96

progeny examined by Southern blotting, none contained a Cre-recombined P2Br allele.

It has also been discovered that a loxP-flanked neo cassette in the glutamate receptor R6 subunit locus is efficiently recombined by ProCre nucleic acid constructs in mice.

Example 4

ProCre Nucleic acid construct Expression is Highly Tissue-Specific

10 Genomic DNAs from ten different tissues of five- to seven-week old males that contained both a ProCre nucleic acid construct and a P2Bc target allele were analyzed in Southern blots. Southern blots were prepared of Pst I digested DNA from testes (T) and one other tissue (K, 15 kidney; B, brain; S, spleen) of males heterozygous for one of four ProCre nucleic acid constructs and the P2Bc allele. Testis DNA from each male shows a P2Br allele signal, in addition to those generated by the wild-type RP2 (WT) and P2Bc alleles. Other tissues show only the WT and P2Bc 20 signals. Only the testis samples showed signal indicating Cre-mediated recombination of the target. The intensity of the P2Br signal relative to that of the wild-type allele ranged from 10% to 22% for different ProCre strains and did not correlate with the ProCre nucleic acid construct copy 25 number. The copy number of ProCre nucleic acid constructs varied among lines showing similar levels of recombination in testis. For example, restriction patterns and densitometric analyses showed that line 58 contained a single copy of the ProCre nucleic acid construct, yet 30 showed virtually the same testis recombination signal as line containing more than 100 copies. This variability is

similar to results obtained with other mP1 promoter-driven nucleic acid constructs (Peschon et al. (1987) *Proc Natl Acad Sci USA* 84:5316-5319; Zambrowicz et al. (1993) *Proc Natl Acad Sci USA* 90:5071-5075).

5 As a more sensitive measure of ectopic recombination, PCR amplifications were performed on the same samples. The amplification primers were expected to produce a 325 bp product from the recombined target and a 1.4 kb fragment from the unrecombined allele (Figure 1). The assay was
10 expected to measure the cumulative level of recombination, for any P2Br alleles formed during transient expression of Cre during development would be preserved and perhaps amplified in descendant cells. Low levels of ectopic recombination product were observed in some tissues of all
15 ProCre lines except for one. A southern blot of PCR amplification products of the P2Br allele utilized tissues from a male heterozygous for the ProCre nucleic acid construct and the P2Bc allele. DNA from 10 different tissues was amplified using primers and conditions that
20 produced a 350 bp product from the recombined, P2Br allele. Each lane contains 10% of the reactions, except for the testis reactions, which were diluted 500 (T5), 250 (T2), and 100 (T1) fold prior to loading, and a liver reconstruction control (C), which was diluted 1:100 before
25 loading. The highest level of ectopic activity was observed in cardiac ventricular muscle of mice; in these samples the ectopic signal was more than 100 fold lower than that observed in testis. Three strains showed much lower levels of recombination in brain tissue, and
30 strain 75 additionally showed ectopic activity in spleen. Despite the difficulty of quantifying PCR results, these data clearly indicate that ectopic activity occurred at very low levels in most tissues of most ProCre lines.

Example 5Isolation of Homologously Recombined ProCre ES Cell
Clones Using Targeting Vectors with a loxP-Flanked
Selectable Marker

5 Four male +/-ProCre ES cell lines were established from
129/Sv strain ProCre transgenic mice. In preliminary
experiments, passage 5 cells from one of these lines (PC3)
were used to generate three male chimeras with between 50
and 95% coat color chimerism. In matings with C57BL/6
10 females, two of these male chimeras have sired a total of
11 pups, all bearing the Agouti coat color signifying
germline transmission of the ES cell genome, and 6 of 9
pups genotyped additionally contained the line 70 ProCre
nucleic acid construct. The frequency of germline
15 transmission has not yet been determined, nor has it been
determined whether competency for germline transmission
will persist in homologously recombined ProCre ES cells at
later passages.

To determine if homologously recombined ProCre ES cell
20 clones could be isolated using targeting vectors that
contained a loxP-flanked selectable marker, two
transfections were done using variants of a targeting
vector in which a loxP-flanked neomycin cassette was
inserted into an Nru I site in the *hoxb-1* locus promoter
25 (Figure 2). A Southern blot of BamHI-digested genomic DNAs
were harvested from a 96-well plate from 10 doubly-selected
ES cell clones and hybridized with a probe (shown in Figure
2) which is external to the targeting construct. All
samples show the 7.5 kb band from the wild-type allele and
30 four clones additionally show the 6 kb band predicted to
result from homologous recombination. In these
transfections, 12 of 62 (19%) PC3- and 10 of 56 (18%)

PCs-derived clones that were ganciclovir and G418-resistant (Mansour et al. (1988) *Nature* 336:348-352) were found to be homologously recombined. In two parallel transfections of CCE cells (Robertson et al. (1986) *Nature* 323:) with the same vectors, 32 of 93 (34%) and 15 of 132 (11%) clones were homologously recombined. The total numbers of G418-resistant clones recovered from ProCre ES cell transfections were reduced relative to the parallel CCE transfections. This may be attributable to both Cre-mediated excision of the neomycin cassette and to the fact that the transfections were done under electroporation conditions optimized for CCE cells.

Because it was formally possible that the homologously recombined clones contained inactive loxP sites, five homologously recombined PC3 ES cell clones and the parental PC3 cell line using the primers shown in Figure 2 were either mock transfected or transiently transfected with the pOG231 Cre expression vector. For the transient transfection assay, DNA was harvested 48 hours after transfection and used in PCR assays to assess whether the loxP sites in the recombinant clones could be recombined by Cre. In all cases a clear recombination signal was observed in the pOG231 transfected sample. The recombinant clones and parental cell lines show the 204 bp amplification product of the wild-type allele, and the recombinant clones additionally show a 1600 bp product (1600) resulting from amplification across the neomycin cassette and a nonspecific 1100 bp amplification product (NS). The pOG231-transfected recombinant clones show an additional 268 bp product signaling the Cre-mediated excision of the neomycin cassette from the recombinant alleles of some cells. Experiments were also done to assess the stability of the loxP-flanked neo cassette in

ProCre ES cells. Five recombinant clones were grown in the presence of G418 for two weeks, and then aliquots of each were grown either in the presence or absence of G418 for a further 10 days. PCR assays were performed to determine if
5 Cre-recombined alleles were present in any of these samples and none was observed in the mock transfected controls. These data suggest that there is not enough Cre activity to significantly influence either the ability to isolate recombinant clones or the stability of the selectable
10 markers in those clones, establishing that the loxP sites in these clones were functional.

To determine if there was any detectable Cre activity in ProCre ES cells, aliquots of two lines (PC3 and PC5) were transiently transfected with the targeting vector used
15 to create the P2Bc allele. DNA was recovered 48 hours after transfection and used for PCR amplifications of the P2Br plasmid molecules that would be generated by extrachromosomal Cre-mediated recombination. Small amounts of recombination product were seen in both ProCre ES cell
20 transfections, and none was observed in parallel samples of CCE ES cells. This shows that the ProCre ES cell lines express sufficient Cre to recombine some extrachromosomal targets when the latter are present at high copy numbers.

Example 6Plant DNA Constructs

To define sequences in the LAT52 and LAT59 promoters involved in expression in pollen, proximal promoters were constructed employing a series of linker substitution mutants using the particle bombardment system (Klein et al. (1987) *Nature* 327:70-73; Twell et al. (1989b) *Plant Physiol* 91:1270-1274). These experiments were performed by co-bombarding the test plasmids (luciferase [LUC]-recombinase fusions) with reference plasmids (β -glucuronidase [GUS] fusions). The latter served as a control for bombardment variability and allowed comparisons to be made between independent bombardments.

The context of the -100 promoter in LAT52 and the -115 promoter in LAT59 was chosen because these promoters appeared to be the minimal regions that still conferred high levels (25% relative to the available full-length promoter) of pollen-specific expression (Twell et al. (1991) *Gen Dev* 5:496-507). These minimal promoters were then fused to the Cre coding sequence operatively linked to the luc gene (Ow et al. (1986) *Science* 234:856-858) coding region, and the resulting plasmids served as a basis for creating the nucleic acid constructs. The LAT52 linker substitutions were performed in p52LUC, which contain entire LAT52 5' untranslated region (5' UTR). A series of six 9- to 10-bp-long linker substitutions were made in p52LUC, spanning the region -84 to -29 (52LS1 to 52LS6).

Example 7Tissue Specificity in Plants

The results obtained by transient expression in pollen and in transgenic plants provided information on the effect of the various constructs on expression in pollen but not on their effect on tissue specificity. A tobacco cell culture, TXD (maintained as described by Howard et al. (1992) Cell 68:109-118), was, therefore, added as an additional component of the transient assay system. The TXD cell culture was initiated from tobacco mesophyll cells and therefore represents somatic tissue, as opposed to the gametophytic tissue represented by pollen. Cells in culture were chosen, rather than intact tissue, as the somatic tissue source because such cells superficially resemble pollen in that they can be spread out as a monolayer on a plate before bombardment.

In this experiment, translation fusions between the *luc* coding region and either the CaMV 35S promoter drove strong expression in cell culture but negligible expression in pollen, whereas the LAT52 promoter showed the opposite pattern of strong activity in pollen and negligible activity in cell culture. Thus, the transient assay system mimics the expression pattern observed for these promoters in transgenic plants (Twell et al. (1991) Genes Dev 5:496-507). This differential expression provided us with a tool with which to address tissue specificity.

Example 8Plant Transformation and Analysis of Transgenic Plants

Constructs cloned into pBin19 were introduced into tomato (*Lycopersicon esculentum* cv VF36) by *Agrobacterium*

tumefaciens LBA4404 as previously described (McCormick (1991b) Transformation of tomato with *Agrobacterium tumefaciens*, In Plant Tissue Culture Manual, K. Linsey, Ed B6:1-9). At least 20 independent transformants were
5 obtained for each construct.

For β -glucuronidase (GUS) assays, 5 to 20 μ L of pollen, pooled from several flowers of the same plant, was ground directly in Eppendorf tubes in 50 to 100 μ L of GUS extraction buffer (Jefferson et al. (1987) *EMBO* 6:3901-
10 3907) using a Teflon-tipped homogenizer driven by a drill. Expression in pollen was measured by fluorometrically assaying GUS activity in supernatants of pollen extracts using 2mM 4-methylumbelliferyl β -D-glucuronide (Sigma) as
15 activity was corrected for variation in total protein content using a bicinchoninic acid protein assay kit (Pierce, Rockford, IL).

Expression in leaves, flowers, stems, roots, and seed was tested histochemically by staining with 5-bromo-4-
20 chloro-3-indolyl β -D-glucuronide (Molecular Probes, Eugene, OR) as described previously (Jefferson et al. (1987) *EMBO* 6:3901-3907). Expression in leaves was also analyzed fluorometrically as given previously.

Example 9

25 Transient Transformation of Tobacco Pollen and Cell Culture

Pollen spread out as a monolayer was bombarded essentially as previously described (Twell et al. (1991) *Genes Dev* 5:496-507), except that gold was substituted for
30 tungsten and only 1 μ g of test plasmid and used per plate.

TXD cell culture (maintained as described by Howard et al. (1992) *Cell* 68:109-118) was spread out similarly as a monolayer (1 mL of a 50-mL stationary culture per plate) and bombarded as previously described. Between six and 12 independent bombardments were performed for each construct. In each experiment, the test plasmid was co-bombarded with a reference plasmid: pB1223 (Clontech, Palo Alto, CA) was used for assays of all constructs in tobacco cell culture; pLAT59-12 (Twell et al. (1990) *Development* 109:705-713) for assays of LAT52 and LAT56 constructs in tobacco pollen; pLAT56-12 (Twell et al. (1990) *Development* 109:705-713) for assays of LAT59 constructs in tobacco pollen. Processing of the tissue after ~ 15 to 17 hr and analysis of GUS and LUC activity were as described previously (Twell et al. (1991) *Genes Dev* 5:496-507). Transient expression was reported as "relative LUC activity," which represents the ratio between the test (LUC) and the reference (GUS) plasmids.

While the invention has been described in detail with reference to certain preferred embodiments thereof, it will be understood that modifications and variations are within the spirit and scope of that which is described and claimed.

That which is claimed is:

1. A nucleic acid construct comprising a germline-specific promoter operatively associated with a recombinase coding sequence.
2. A nucleic acid construct according to claim 1 wherein said germline-specific promoter is the protamine 1 gene promoter, the protamine 2 gene promoter, the spermatid-specific promoter from the c-kit gene, the sperm-specific promoter from angiotensin-converting enzyme, oocyte specific promoter from the ZP1 gene, oocyte specific promoter from the ZP2 gene, or oocyte specific promoter from the ZP3 gene.
3. A nucleic acid construct according to claim 1 wherein said germline-specific promoter is the LAT52 gene promoter from tomato, the LAT56 gene promoter from tomato, the LAT59 gene promoter from tomato, the pollen-specific promoter of the Brassica S locus glycoprotein gene, or the pollen-specific promoter of the NTP303 gene.
4. A nucleic acid construct according to claim 1 wherein said recombinase coding sequence encodes Cre recombinase.
5. A nucleic acid construct according to claim 4 wherein said construct is ProCre, comprising the protamine 1 gene promoter operatively associated with Cre recombinase.

6. A nucleic acid construct according to claim 1 wherein said recombinase coding sequence encodes FLP recombinase.

7. A nucleic acid construct according to claim 6 wherein said construct is ProFLP, comprising the protamine 1 gene promoter operatively associated with FLP recombinase.

8. A nucleic acid construct according to claim 1 wherein said recombinase coding sequence encodes the R gene product of *Zygosaccharomyces*.

9. A nucleic acid construct according to claim 8 wherein said construct is ProR, comprising the protamine 1 gene promoter operatively associated with the R gene product of *Zygosaccharomyces*.

10. A nucleic acid construct comprising a conditional promoter operatively associated with a recombinase coding sequence.

11. A nucleic acid construct comprising a tissue-specific promoter operatively associated with a recombinase coding sequence.

12. Embryonic stem cells containing a nucleic acid construct according to claim 1.

13. Embryonic stem cells according to claim 12 wherein the genome thereof comprises a transcriptionally

active selectable marker flanked by two recombination target sites.

14. Embryonic stem cells according to claim 13 wherein the recombinase encoded by the recombinase coding sequence operatively associated with a germline-specific promoter is selective for the recombination target sites flanking said selectable marker.

15. Embryonic stem cells according to claim 13 further comprising one or more of:

a nucleic acid fragment flanked by two recombination target sites, wherein said recombination target sites are different than the recombination target sites which flank said selectable marker,

a nucleic acid construct comprising a conditional promoter operatively associated with a recombinase coding sequence, or

a nucleic acid construct comprising a tissue-specific promoter operatively associated with a recombinase coding sequence.

16. Embryonic stem cells containing a nucleic acid construct according to claim 2.

17. Embryonic stem cells containing a nucleic acid construct according to claim 3.

18. Embryonic stem cells containing a nucleic acid construct according to claim 4.

19. Embryonic stem cells containing a nucleic acid construct according to claim 5.

20. Embryonic stem cells containing a nucleic acid construct according to claim 6.

21. Embryonic stem cells containing a nucleic acid construct according to claim 7.

22. Embryonic stem cells containing a nucleic acid construct according to claim 8.

23. Embryonic stem cells containing a nucleic acid construct according to claim 9.

24. Embryonic stem cells containing a nucleic acid construct according to claim 10.

25. Embryonic stem cells according to claim 24 wherein the genome thereof comprises a transcriptionally active selectable marker flanked by two recombination target sites.

26. Embryonic stem cells containing a nucleic acid construct according to claim 11.

27. Embryonic stem cells according to claim 26 wherein the genome thereof comprises a transcriptionally active selectable marker flanked by two recombination target sites.

28. A method for excission of the transcriptionally active selectable marker from the embryonic stem cells of claim 13, said method comprising:

passaging the genome derived from said embryonic stem cells through gametogenesis.

29. A method according to claim 28 wherein said genome is passaged through spermatogenesis.

30. A method according to claim 28 wherein said genome is passaged through oogenesis.

31. A method according to claim 28 wherein said embryonic stem cells further comprise one or more of:

a nucleic acid fragment flanked by two recombination target sites, wherein said recombination target sites are different than the recombination target sites which flank said selectable marker,

a nucleic acid construct comprising a conditional promoter operatively associated with a recombinase coding sequence, or

a nucleic acid construct comprising a tissue-specific promoter operatively associated with a recombinase coding sequence.

32. A method for the production of recombinant alleles, said method comprising:

introducing a nucleic acid fragment flanked by at least two recombination target sites into embryonic stem cells according to claim 10, and

passaging the genome derived from said embryonic stem cells through gametogenesis.

33. A method according to claim 32 wherein said nucleic acid fragment comprises an essential portion of a gene of interest.

34. A method according to claim 32 wherein said nucleic acid fragment is introduced by homologous recombination, random insertion, retroviral insertion, or site specific-mediated recombination.

35. A method for the production of recombinant alleles, said method comprising:

introducing a nucleic acid fragment flanked by at least two recombination target sites into embryonic stem cells according to claim 13, and

passaging the genome derived from said embryonic stem cells through gametogenesis.

36. A method according to claim 35 wherein said embryonic stem cells further comprise a second nucleic acid construct selected from the group consisting of a construct comprising a conditional promoter operatively associated with a recombinase coding sequence and a construct comprising a tissue-specific promoter operatively associated with a recombinase coding sequence.

37. A method according to claim 36 wherein the recombinase encoded by said second construct is expressed in response to inducing conditions.

38. A method according to claim 36 wherein the recombinase encoded by said second construct is expressed in a tissue selective manner.

39. A method according to claim 35 wherein the recombination target sites flanking said nucleic acid fragment are recognized by a recombinase which is expressed under the control of a conditional promoter or a tissue specific promoter.

40. A method for the production of recombinant alleles, said method comprising:

introducing at least one recombinase responsive construct into embryonic stem cells according to claim 10,
wherein said construct(s) comprise(s) a nucleic acid fragment and a selectable marker,
wherein said selectable marker is flanked by a first pair of recombination target sites, and
wherein said nucleic acid fragment is flanked by a second pair of recombination target sites,

passaging the genome derived from said embryonic stem cells through gametogenesis.

41. A method according to claim 40 wherein said first pair of recombination target sites is recognized by a recombinase which is expressed under the control of a germline-specific promoter and said second pair of recombination target sites is recognized by a recombinase.

which is expressed under the control of a conditional promoter or a tissue specific promoter.

42. A method according to claim 40 wherein said embryonic stem cells further comprise a second nucleic acid construct selected from the group consisting of a construct comprising a conditional promoter operatively associated with a recombinase coding sequence and a construct comprising a tissue-specific promoter operatively associated with a recombinase coding sequence.

43. A method for the conditional assembly of functional gene(s) for expression in eukaryotic cells by recombination of individual inactive gene segments from one or more gene(s) of interest,

wherein each of said segments contains at least one recombination target site, and

wherein at least one of said segments contains at least two recombination target sites,

said method comprising:

introducing said individual inactive gene segments into an embryonic stem cell according to claim 10, thereby providing a DNA which encodes a functional gene of interest, the expression product of which is biologically active, upon passage of the genome derived from said stem cells through gametogenesis.

44. A method for the generation of recombinant livestock, said method comprising:

combining embryonic stem cells that include a nucleic acid construct according to claim 1 with host pluripotential ES cells derived from early preimplantation embryos, and

introducing these combined embryos into a host female and

allowing the derived embryos to come to term.

45. A method for the generation of recombinant plants, said method comprising transforming plant zygotes with nucleic acid constructs according to claim 1 and allowing the zygote to develop.

1/2

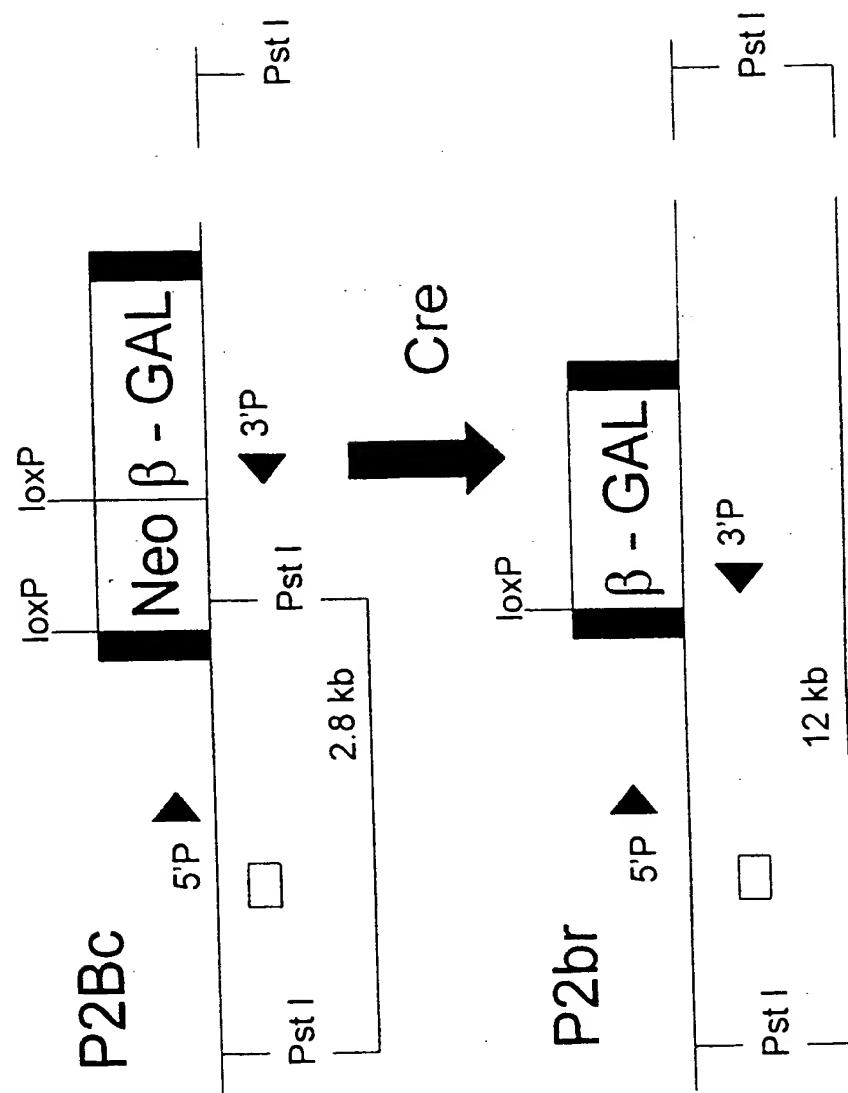


Figure 1

2/2

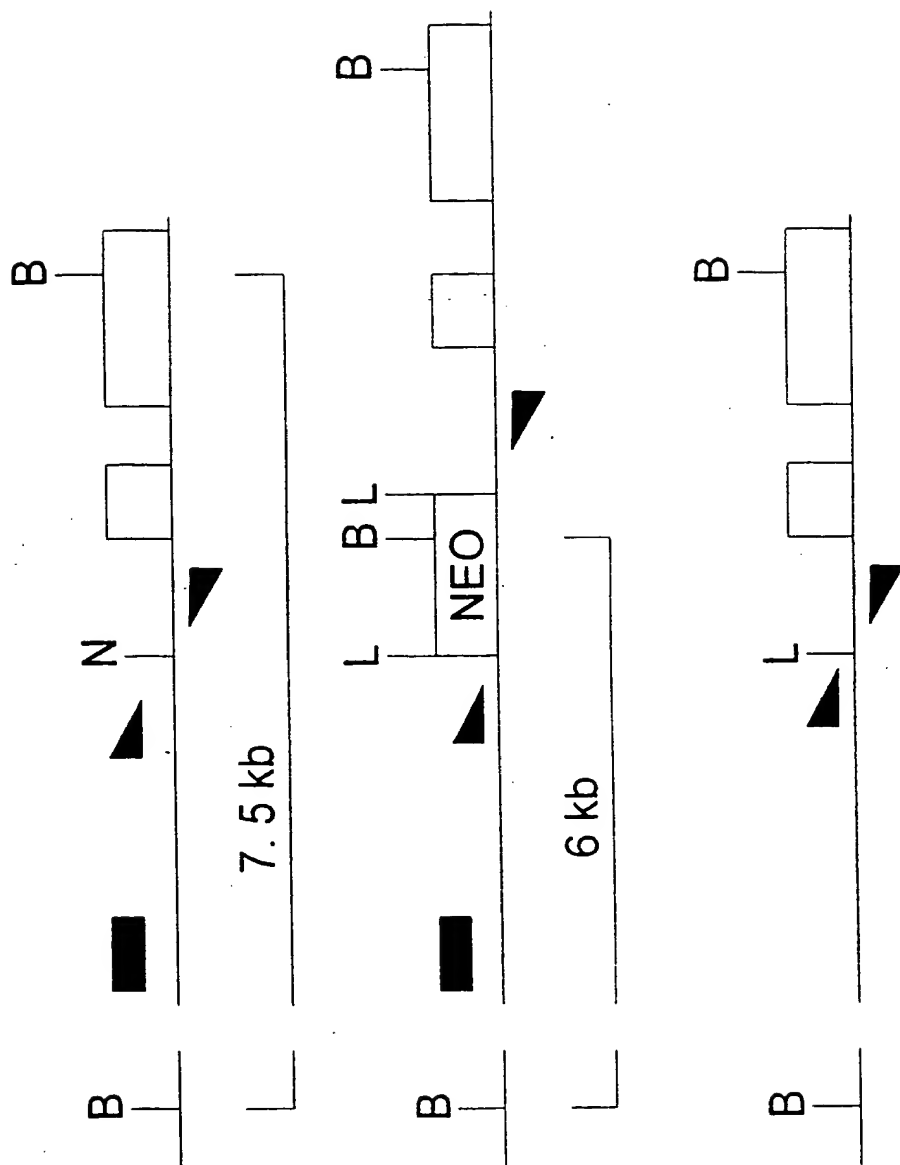


Figure 2

SEQUENCE LISTING

<110> O'Gorman, Steve
Wahl, Geoffrey

<120> Site-Specific Germline Recombination in
Eukaryotes and Constructs Useful Therefor

<130> Salk2190

<150> 08/919,501

<151> 1997-08-28

<160> 8

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 652

<212> DNA

<213> Mus musculus

<400> 1

gtctagtaat gtccaacacc tccctcagtc caaacactgc tctgcatcca tgtggctccc	60
atttatacct gaagcacttg atggggcctc aatgttttac tagagccac cccctgcaa	120
ctctgagacc ctctggattt gtctgtcagt gcctcactgg ggcgttgat aatttcttaa	180
aagggtcaagt tccctcagca gcattctctg agcagtctga agatgtgtgc ttccacagtt	240
acaaatccat gtggctgttt caccacactg cctggccttg ggttatctat caggacctag	300
cctagaagca ggtgtgtggc acttaacacc taagctgagt gactaactga aactcaagt	360
ggatgccatc ttgtcactt cttgactgtg acacaagcaa ctctgatgc caaagccctg	420
cccacccctc tcatgcccat atttgacat ggtacaggtc ctactggcc atggctctgtg	480
aggtcctggg cctctttgac ttcataattc ctaggggcca ctagtatcta taagaggaag	540
aggggtgctgg ctcccaggcc acagcccaca aaattccacc tgctcacagg ttggctggct	600
cgaccagggt ggtgtccctt gctctgagcc agctcccggc caagccagca cc	652

<210> 2

<211> 29

<212> DNA

<213> Artificial Sequence

<400> 2

gtctagtaat gtccaacacc tccctcagt	29
---------------------------------	----

<210> 3

<211> 31

<212> DNA

<213> Artificial Sequence

<400> 3

ctctgagcca gctcccggcc aagccagcac c	31
------------------------------------	----

<210> 4

<211> 1022

<212> DNA

<213> Artificial Sequence

<400> 4

atggagcaaa agctgatttc tgaggaggat ctgggaggac ccaagaagaa gaggaagggtg	60
tccaatttac tgaccgtaca ccaaaatttg cctgcattac cggtcgatgc aacgagtgat	120
gagggttcgca agaacctgat ggacatgttc agggatcgcc aggcgttttc tgagcatacc	180

tggaataatgc	ttctgtccgt	ttgccggctcg	tggggggcat	ggtgcaagt	aataaccgga	240
aatgggtttcc	cgcagaacct	gaagatgttc	gcgattatct	tctatatctt	caggcgcgcg	300
gtctggcagt	aaaaactatc	cagcaacatt	tggggccagct	aaacatgctt	catcgctcgt	360
ccgggctgcc	acgaccaagt	gacagcaatg	ctgttttact	ggttatgcgg	cggatccgaa	420
aagaaaacgt	tgatgccggg	gaacgtgcaa	aacaggctct	agcgttcgaa	cgcactgatt	480
tcgaccaggt	tcgttctact	atggaaaata	gcgatcgctg	ccaggatata	cgtaattctg	540
catttctggg	gattgcttat	aacacctgt	tacgtatagc	cgaaattgcc	aggatcagg	600
ttaaagatat	ctcacgtact	gacggtgagg	gaatgttaat	ccatattggc	agaacgaaaa	660
cgctgggttag	caccgcaggt	gtagagaagg	cacttagcct	gggggtaact	aaactggctc	720
agcgatggat	ttccgtctct	ggtgtagctg	atgatccgaa	taactacctg	ttttgccggg	780
tcagaaaaaa	tggtgttgcc	gcgccatctg	ccaccagcca	gctatcaact	cgcgccctgg	840
aagggtatctt	tgaagcaact	catcgattga	tttacggcgc	taaggatgac	tctggtcaga	900
gatacctggc	ctgggtctgga	cacagtgcgc	gtgtcggagc	cgcgcgagat	atggcccgcg	960
ctggagtttc	aataaccggag	atcatgcaag	ctggtggctg	gaccaatgta	aatattgtca	1020
tg						1022

<210> 5

<211> 2293

<212> DNA

<213> Artificial Sequence

<400> 5

gtctagtaat	gtccaacacc	tccctcagtc	caaacactgc	tctgcatcca	tgtggctccc	60
atttataacct	gaagcacttg	atggggcctc	aatgttttac	tagagccac	ccccctgcaa	120
ctctgagacc	ctctggattt	gtctgtcagt	gcctcactgg	ggcggtggat	aatttcttaa	180
aaggtaagt	tccctcagca	gcattctctg	agcagtcgta	agatgtgtgc	tttcacagtt	240
acaaatccat	gtggctgttt	cacccacctg	cctggccttg	ggttatctat	caggacctag	300
cctagaagca	ggtgtgtggc	acttaacacc	taagctgagt	gactaactga	acactcaagt	360
ggatgccatc	tttgtcactt	cttgactgtg	acacaagcaa	ctcctgatgc	caaagccctg	420
cccacccctc	tcattgccat	atttggacat	ggtacaggtc	ctcactggcc	atggtctgtg	480
aggctcctgg	cctctttgac	ttcataattc	ctagggggcca	ctagtatcta	taagagggaag	540
aggggtgctg	ctcccaggcc	acagcccaca	aaattccacc	tgtcacagg	ttggctggct	600
cgacccaggt	ggtgtccctt	gctctgagcc	agctcccggc	caagccagca	cccgggacca	660
tggagcaaaa	gctgatttct	gaggaggatc	tgggaggacc	caagaagaag	aggaagggtg	720
ccaattttact	gaccgtacac	caaaatttgc	ctgcattacc	ggctgatgca	acgagtgtatg	780
aggttcgcaa	gaacctgatg	gacatgttca	gggatcgcca	ggcgttttct	gagcatacct	840
ggaaaatgct	tctgtccgtt	tgccggctcg	ggggggcatg	gtgcaagtgt	aataaccgga	900
aatgggtttcc	cgcagaacct	gaagatgttc	gcgattatct	tctatatctt	caggcgcgcg	960
gtctggcagt	aaaaactatc	cagcaacatt	tggggccagct	aaacatgctt	catcgctcgt	1020
ccgggctgcc	acgaccaagt	gacagcaatg	ctgttttact	ggttatgcgg	cggatccgaa	1080
aagaaaacgt	tgatgccggg	gaacgtgcaa	aacaggctct	agcgttcgaa	cgcactgatt	1140
tcgaccaggt	tcgttctact	atggaaaata	gcgatcgctg	ccaggatata	cgtaattctg	1200
catttctggg	gattgcttat	aacacctgt	tacgtatagc	cgaaattgcc	aggatcagg	1260
ttaaagatat	ctcacgtact	gacggtgagg	gaatgttaat	ccatattggc	agaacgaaaa	1320
cgctgggttag	caccgcaggt	gtagagaagg	cacttagcct	gggggtaact	aaactggctc	1380
agcgatggat	ttccgtctct	ggtgtagctg	atgatccgaa	taactacctg	ttttgccggg	1440
tcagaaaaaa	tggtgttgcc	gcgccatctg	ccaccagcca	gctatcaact	cgcgccctgg	1500
aagggtatctt	tgaagcaact	catcgattga	tttacggcgc	taaggatgac	tctggtcaga	1560
gatacctggc	ctgggtctgga	cacagtgcgc	gtgtcggagc	cgcgcgagat	atggcccgcg	1620
ctggagtttc	aataccggag	atcatgcaag	ctggtggctg	gaccaatgta	aatattgtca	1680
tgaactatat	ccgtaacctg	gatagtgaag	caggggcaat	ggtgcgcctg	ctggaagatg	1740
gcgattagcc	attaacgcgt	aaatgattgc	tataattatt	tgatatttat	ggtgacatat	1800
gagaaaggat	ttcaacatcg	acggaaaata	tgtagtgtcg	tctgtaagca	ctaattattca	1860
gtcgccagcc	gacattgtca	ctgtaaagct	gagcgataga	atgcctgata	ttgactcaat	1920
atccgggtgcg	tttctgttca	aaagtatgcg	tagtgctgaa	catttcgcga	tgaatcccac	1980
cgaggaagaa	gcacggcgcg	gttttgetaa	agtgtgtctt	gagtttggcg	aactcttggg	2040
taagggttga	attgtcgagg	ctgggtgttg	cggaccgcta	tcaggacata	gcgttggcta	2100
cccgtgatat	tgctgaagag	cttggcgggc	aatgggctga	ccgttctctc	gtgctttacg	2160
gtatcgccgc	tcccgattcg	cagcgcatcg	ccttctatcg	ccttcttgac	gagttcttct	2220
gaggggatcg	gcaataaaaa	gacagaataa	aacgcacggg	tggtgggtcg	tttgttcgga	2280
tcgatccgtc	gac					2293

<210> 6
 <211> 86
 <212> DNA
 <213> Artificial Sequence

<400> 6
 cccgggatca attcaccatg ggaataaactt cgtatagcat acattatacg aagttatgga 60
 tccgccgcta tcaggacata gcgttg 86

<210> 7
 <211> 4172
 <212> DNA
 <213> Artificial Sequence

<400> 7
 gcacttttctg gggaaatgtg cgcggaaccc ctatttggtt atttttctaa atacattcaa 60
 atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat tgaaaaagga 120
 agagtatgag tattcaacat ttccgtgtcg cccttattcc cttttttgctg gcattttgccc 180
 ttctgtgttt tgctcaccga gaaacgctgg tgaagtaaa agatgctgaa gatcagttgg 240
 gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt gagagttttc 300
 gcccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt ggcgcggtat 360
 tatcccgat tgacgcccgg caagagcaac tcggtcgccg catacactat tctcagaatg 420
 acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag 480
 aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta cttctgacaa 540
 cgatcgaggg accgaaggag ctaaccgctt ttttgacaaa catgggggat catgtaactc 600
 gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca 660
 cgatgcctgt agcaatggca acaacgttgc gcaaacatt aactggcgaa ctacttactc 720
 tagcttccc gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc 780
 tgctctcgcc ccttcgggct ggctggttta ttgctgataa atctggagcc ggtgagcgtg 840
 ggtctcgccg tatcattgca gcactggggc cagatggtaa gccctcccgt atcgtagtta 900
 tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag 960
 gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat atactttaga 1020
 ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt tttgataatc 1080
 tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa 1140
 agatcaaagg atcttcttga gatccttttt tctcgccgt aatctgctgc ttgcaaaaca 1200
 aaaaaccacc gctaccagcg gtggtttggt tgccggatca agagctacca actcttttc 1260
 cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgtccttcta gtgtagccgt 1320
 agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatac 1380
 tgttaccagt ggtgctgccc agtggcgata agtcgtgtct taccgggttg gactcaagac 1440
 gatagttacc ggataaggcg cagcggctcg gctgaacggg ggggttcgtgc acacagccca 1500
 gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaaagc 1560
 ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag 1620
 gagagcgcac gagggagctt ccagggggaa acgcctggtt tctttatagt cctgtcgggt 1680
 ttccgccact ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat 1740
 ggaaaaacgc cagcaacgcg gcctttttac ggttccctggc cttttgctgg ccttttctc 1800
 acatgttctt tccctcggtt tcccttgatt ctgtggataa ccgtattacc gcctttgagt 1860
 gagctgatac cgctcgccgc agccgaacga ccgagcgag cgagtcagtg agcgaggag 1920
 cggagagcgc cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatgca 1980
 gctggcacga caggtttccc gactggaaag cgggcagtga gcgcaacgca attaatgtga 2040
 gttagctcac tcattaggca cccaggtt tacactttat gcttccggt cgtagtgtgt 2100
 gtggaattgt gagcgataa caatttcaca caggaaacag ctatgacctat gattacgcca 2160
 agctcgaaat taaccctcac taaagggaac aaaagctggg tacgaattca gatctccgg 2220
 gatcaattca ccatgggaat aacttcgtat agcatacatt atacgaagtt atggatccgg 2280
 tcgagcagtg tggttttgca agaggaagca aaaagcctct ccaccaggg ctggaatggt 2340
 tccacccaat gtcgagcagt gtggttttgc aagaggaagc aaaaagcctc tccaccagg 2400
 cctggaatgt ttccacccaa tgcgagcaaa accccgcccc gcgtcttctc attggcgaat 2460
 tcgaacacgc agatgcagtc gggggcgccg ggtcccaggt ccacttcgca tattaagggtg 2520
 acgcgtgtgg cctcgaacac cgagcgaccc tgcagccaat atgggatcgg ccattgaaca 2580
 agatggattg cacgcagggt ctccggccgc ttgggtggag aggtatttcg gctatgactg 2640
 ggcacaacag acaatcggct gctctgatgc cgccgtgttc cggctgtcag cgcagggggc 2700
 cccggttctt tttgtcaaga ccgacctgtc cggtgccctg aatgaactgc aggacgaggc 2760
 agcgcggtcta tcgtggcttg ccacgacggg cgttccttgc gcagctgtgc tcgacgttgt 2820

cactgaagcg	ggaagggact	ggctgctatt	gggcgaagtg	ccggggcagg	atctcctgtc	2880
atctcacctt	gtccctgccg	agaaagtatc	catcatggct	gatgcaatgc	ggcggctgca	2940
tacgcttgat	ccggctacct	gcccattcga	ccaccaagcg	aaacatcgca	tcgagcgagc	3000
acgtactcgg	atggaagccg	gtcttgctga	tcaggatgat	ctggacgaag	agcatcaggg	3060
gctcgcgcc	gccgaactgt	tcgccaggct	caaggcgcgc	atgcccgacg	gcgaggatct	3120
cgtcgtgacc	catggcgatg	cctgcttgcc	gaatatcatg	gtggaaaatg	gccgcttttc	3180
tggattcatc	gactgtggcc	ggctgggtgt	ggcggaccgc	tatcaggaca	tagcgttggc	3240
taccctgat	attgctgaag	agcttgccg	cgaatgggct	gaccgcttcc	tcgtgcttta	3300
cggatcgcgc	gctcccgatt	cgcagcgc	cgccttctat	cgccttcttg	acgagttctt	3360
ctgaggggat	cggcaataaa	aagacagaat	aaaacgcacg	ggtgttgggt	cgtttgttcg	3420
gatagggatc	aattcaccat	gggaataact	tcgtatagca	tacattatac	gaagttatgg	3480
atccactagt	tctagagcgg	ccgccaccgc	ggtggagctc	caattcgccc	tatagtga	3540
cgtattacaa	ttcactggcc	gtcgttttac	aacgtcgtga	ctgggaaaac	cctggcgcta	3600
cccaacttaa	tcgccttgca	gcacatcccc	ctttcgccag	ctggcgtaat	agcgaagagg	3660
cccgaccga	tcgccttcc	caacagttgc	gcagcctgaa	tggcgaatgg	gacgcgccct	3720
gtagcggcgc	attaagcgcg	gcgggtgtgg	tggttacgcg	cagcgtgacc	gctacacttg	3780
ccagcgcct	agcgcgcgt	ccttccgctt	tcttcccttc	ctttctcgcc	acgttcgccc	3840
gctttccccg	tcaagctcta	aatcgggggc	tccctttagg	gttccgattt	agtgccttac	3900
ggcacctcga	ccccaaaaaa	cttgattagg	gtgatggttc	acgtagtggg	ccatcgccct	3960
gatagacggg	ttttcgccct	ttgacgttgg	agtccacgtt	ctttaatagt	ggactcttgt	4020
tccaaactgg	aacaacactc	aaccctatct	cggctctatc	ttttgatcta	taagggattt	4080
tgccgatttc	ggcctattgg	ttaaaaaatg	agctgattta	acaaaaattt	aacgcgaatt	4140
ttaacaaaat	attaacgctt	acaatttagg	tg			4172

<210> 8

<211> 34

<212> DNA

<213> Artificial Sequence

<400> 8

ataaacttcgt atagcataca ttatacgaag ttat

34